Process design and modeling for the production of triacylglycerols (TAGs) in *Rhodococcus opacus* PD630

by

Neidi Miller

B.S. Chemical Engineering
University of Puerto Rico-Mayaguez, 2005

SUBMITTED TO THE DEPARTMENT OF CHEMICAL ENGINEERING IN PARTIAL FULLFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE IN CHEMICAL ENGINEERING
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

FEBRUARY 2012

Signature of Author: _____

Neidi Miller
Department of Chemical Engineering
January 20, 2012

Certified by: _____

Kristala L. Jones Prather
Associate Professor of Chemical Engineering
Thesis Supervisor

Accepted by: _____

William M. Deen
Professor of Chemical Engineering
Chairman, Committee for Graduate Students

Process design and modeling for the production of triacylglycerols (TAGs) in *Rhodococcus opacus* PD630

by

Neidi Miller

Submitted to the Department of Chemical Engineering on January 20, 2012 in Partial Fulfillment of the Requirements for the Degree of Master of Science in Chemical Engineering

ABSTRACT

The oleaginous microorganism *Rhodococcus opacus* PD630 was used to study the characteristics and kinetics of the accumulation of triacylglycerols (TAGs) in cells. In this process, accumulation of TAG is stimulated when a carbon source is present in the medium in excess and the nitrogen source is limiting growth. Under controlled fermentation conditions the organism *Rhodococcus opacus PD630* has been shown to grow to high cell density, producing high yields of TAGs (above 50% of cell dry weight) in a relatively short period of time. In this study, the reaction stoichiometry was established and the carbon balance for the process has been effectively closed, accounting for approximately 91% of the total carbon in the system. Several fed-batch strategies were explored at the 1L benchtop bioreactor scale. Feeding both carbon and ammonium sulfate as the nitrogen source can sustain cell growth but was found to significantly obstruct the accumulation of TAGs. While these fed-batch strategies did not lead to titer improvements, they did highlight the significance of TAG degradation for growth. To aid in future process design strategy optimization an unstructured kinetic model was developed to describe the dynamics of the fermentation of *Rhodococcus opacus* PD630 and its triacylglycerol (TAG) production. The kinetic parameters for this model were either measured from experimental data or estimated by fitting the experimental data using least-squares non-linear regression. Global minimum of the sum of squared errors (SSE) between the model prediction and various experimental data sets was found by an iterative process of parameter space exploration. The minimum SSE obtained was 91.229. The proposed model is the first step towards understanding and optimizing the process of lipid production and accumulation in oleaginous organisms.

Thesis Supervisor: Kristala L. Jones Prather
Title: Associate Professor of Chemical Engineering

DEDICATION

This thesis is dedicated to my loving husband, Stuart Miller, with my deepest expression of love and appreciation for the encouragement and support you gave me and for all the sacrifices you made to make this possible. You are everything to me, thank you for saving my life, in more ways than one.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS AND NOMENCLATURE

ER	endoplasmic reticulum
FAME	fatty acid methyl ester
WS	wax ester synthase
DGAT	diacylglycerol acyltransferase
TAG	triacyglycerol
tFA	total fatty acids
PHA	polyhydroxyalkanoic acid
PL	phospholipid
SLD	small lipid droplets
SSE	sum of squares of error
SOP	standard operating procedure
ODE	ordinary differential equation
NaOH	sodium hydroxide
$(NH_4)_2SO_4$	ammonium sulfate
X	residual biomass concentration, equal to total measured biomass minus measured TAGs, g biomass L^{-1}
N	ammonium sulfate concentration, g $(NH_4)_2SO_4$ L^{-1}
G	glucose concentration, g glucose L^{-1}
P	product concentration, g TAG L^{-1}
α	growth associated TAG production constant, g TAG g biomass^{-1}
β	specific TAG production rate, g TAG g glucose^{-1} h^{-1}
β_{max}	maximum specific TAG production rate, g TAG g biomass^{-1} h^{-1}
μ	specific growth rate, h^{-1}
$\mu_{max,S}$	maximum specific growth rate for growth on substrates, h^{-1}
$\mu_{max,P}$	maximum specific growth rate for growth on product, h^{-1}
K	lumped Michaelis-Menten type constant for substrates, g glucose g $(NH_4)_2SO_4$ L^{-2}
K_G	Michaelis-Menten type constant for glucose, g glucose L^{-1}
K_P	Michaelis-Menten type constant for product, g TAG L^{-1}

$Y_{X/N}$ stoichiometric yield coefficient of residual biomass on nitrogen, g biomass g $(NH_4)_2SO_4^{-1}$

$Y_{X/G}$ stoichiometric yield coefficient of residual biomass on glucose, g biomass g glucose^{-1}

$Y_{X/P}$ stoichiometric yield coefficient of residual biomass on product, g biomass g TAG^{-1}

$Y_{P/G}$ stoichiometric yield coefficient of product on glucose, g TAG g glucose^{-1}

CHAPTER 1
INTRODUCTION

The search for renewable fuels that can potentially reduce or replace our consumption of fossil fuels has intensified significantly in recent years. The biggest limitation to large-scale development and commercialization of renewable fuels is the lack of inexpensive oil feedstocks. Using vegetable oils as the source of triacylglycerol results in high production costs; feedstock costs account for 85% of total production costs of biodiesel (Canakci & Sanli, 2008). To become an economically viable alternative fuel, biodiesel must compete economically with petroleum-based diesel fuel. However, the raw material cost of biodiesel is already higher than the final cost of diesel fuel. Another important consideration is that even if biodiesel remained economically competitive, the current limited worldwide supply of plant oils prevents biodiesel from replacing conventional diesel (Durret et al, 2008). In 2007 the US Department of Agriculture and US Department of Energy estimated that converting the entire 2005 USA soybean crop to biodiesel would replace only 10% of conventional diesel consumed. This has motivated researchers to consider other sources of TAGs, among them, microbial oils, which are produced by some oleaginous microorganisms like yeast, fungi, bacteria and microalgae. Compared to other plant oils, microbial oils have many advantages, such as short life cycle, less labor required, less sensitivity to venue, season or climate and overall simplicity to scale up (Li et al, 2008). Therefore, microbial oils might become a potential oil feedstock for biodiesel production in the future, though there is much research that needs to be carried out to advance this option.

The need for high energy molecules derived from renewable sources has motivated the search for oleaginous organisms that produce microbial oils (Durret et al, 2008; Li et al, 2008; Elbahloul and Steinbuchel, 2010; Kurosawa et al, 2010; Kosa and Ragauskas, 2010). Many bacteria are able to accumulate specialized lipids, such as poly(3-hydroxybutyric acid) or other polyhydroxyalkanoic acids (PHAs). But only bacteria belonging to the actinomycetes group including *Mycobaterium, Rhodococcus, Nocardia* and *Streptomyces* accumulate large amounts of TAGs which serve as storage reservoirs for energy and carbon (Waltermann et al, 2005). Certain *Rhodococcus* species are also able to synthesize and accumulate PHAs after

cultivation on different carbon sources under nitrogen-limiting conditions. One of the species, the *Rhodococcus opacus* strain PD630, is of particular interest because it has been reported to accumulate between 76% to 87% of dry cell weight as acylglycerols when grown on gluconate under nitrogen limiting conditions but is apparently unable to synthesize PHAs (Waltermann et al, 2000; Alvarez & Steinbuchel, 2002). One example of microbial oils is triacylglycerols (TAGs) which are non-polar, water-insoluble triesters of glycerol with fatty acids (Steinbuchel and Alvarez, 2002). In most bacteria, accumulation of TAG and other neutral lipids is usually stimulated if a carbon source is present in the medium in excess and if the nitrogen source is limiting growth. Cellular growth is impaired under these conditions and the cells use the carbon source mainly for the biosynthesis of neutral lipids, therefore the accumulation of TAGs occurs predominantly during stationary phase (Alvarez & Steinbuchel, 2002; Murphy & Vance, 1999).

Biosynthesis of TAG can be subdivided into three steps: (1) production of fatty-acyl-compounds; (2) formation of glycerol intermediates; and (3) sequential esterification of the glycerol moiety with fatty acyl-residue. Figure 1 summarizes the metabolic reactions of key enzymes involved with TAG biosynthesis. Although the TAGs from *R. opacus* vary in composition depending on the carbon source, the stereospecific distribution of the acyl residues on the glycerol backbone is not random. The shorter and saturated fatty acid residues are predominantly esterified to the sn-2 hydroxyl group, whereas unsaturated fatty acids are predominantly bound at position sn-3 (Waltermann et al, 2000).

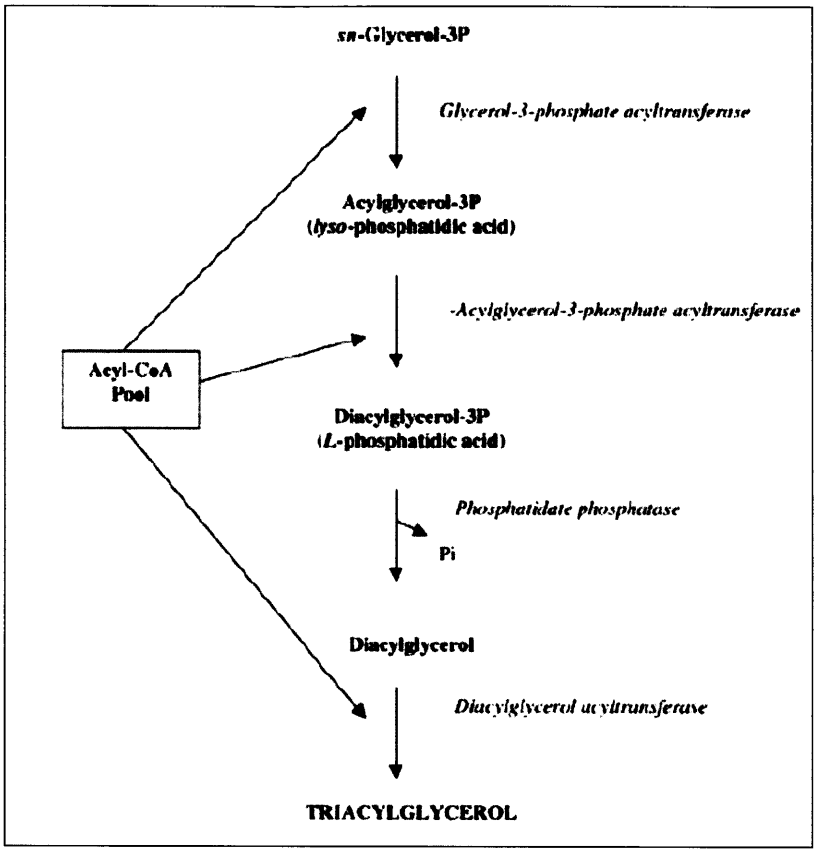

Figure 1: Metabolic reactions of key enzymes involved in the biosynthesis of triacylglycerols (TAGs) and their acylglycerol precursors. This figure was reproduced from Alvarez and Steinbuchel, 2002.

The TAGs are stored in spherical lipid bodies, or cytoplasmic inclusions virtually surrounded by a thin boundary layer. Similar to the formation of eukaryotic lipid bodies at the endoplasmic reticulum (ER) membrane, bacterial neutral body synthesis is strictly associated with the plasma membrane, with one main difference: the bacterial lipids are not synthesized between the leaflets of a phospholipid bilayer (Waltermann & Steinbuchel, 2005). In bacteria, neutral lipid-body formation starts with attachment of wax ester synthase/diacylglycerol acyltransferase (WS/DGAT) to the cytoplasm membrane and subsequent synthesis of small lipid droplets (SLDs) forming an oleogenous layer, which is coated by a phospholipid (PL) monolayer. Lipid-prebodies are formed by conglomeration and coalescence of SLDs leading to the formation of membrane bound lipid-prebodies that are subsequently released and become cytoplasmic lipid-bodies (Waltermann & Steinbuchel, 2005).

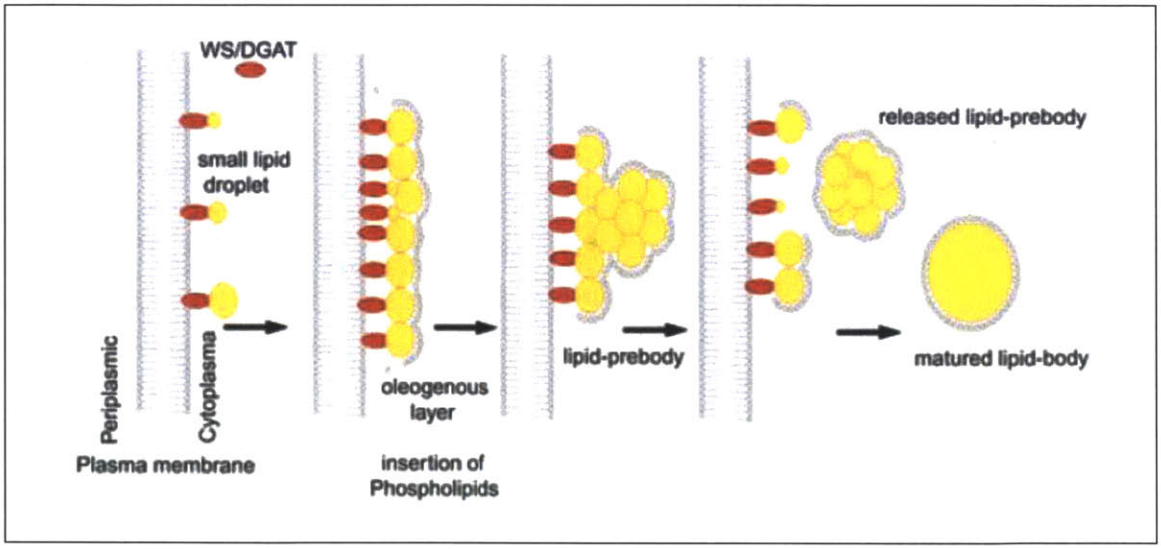

Figure 2: Model for lipid-body formation in prokaryotes. This figure was reproduced from Waltermann & Steinbuchel, 2005.

Under controlled fermentation conditions the organism *Rhodococcus opacus PD630* can grow to very high density, producing high yields of TAGs in a relatively short period of time (Alvarez, 1996; Kurosawa et al, 2010). In order to exploit the potential of *Rhodococcus opacus* as a producer of microbial oils, it will be essential to understand the kinetics of growth and TAG accumulation. Here we report the development of a structured model that describes such kinetics in minimal defined media over a range of substrate ratios. This model should provide a clearer understanding of the TAG accumulation process and lead to improved process designs that can take full advantage of the catalytic capabilities of the strain to maximize lipid production.

CHAPTER 2
MATERIALS AND METHODS

2.1 Microorganism and Medium:

Rhodococcus opacus PD630 (DSM 44193) was used for the production of TAGs. The culture medium used was a phosphate buffered defined medium which contained (per liter): 16.0 g glucose, 1.0g ammonium sulfate $(NH_4)_2SO_4$, 1.0 g magnesium sulfate ($MgSO_4 \cdot 7H_2O$), 0.015 g calcium chloride ($CaCl_2 \cdot 2H_2O$), 1.0mL trace element solution, 1.0 mL stock A solution, and 35.2mL 1.0 M phosphate buffer containing both monobasic and dibasic potassium phosphate. The trace element solution, stock A solution and phosphate buffer were the same described by Chartrain et al (1998). Modifications to the medium's glucose and ammonium sulfate concentration are specified in the discussion below as needed.

2.2 Shake Flask Cultures:

The shake flask experiments were conducted in triplicate using 250 mL baffled flasks containing 100 mL of defined medium incubated on a rotary shaker (250 rpm) at 30°C. The cultures were inoculated with a 3% inoculum of *R. opacus* frozen culture stock. The frozen culture stock was created by mixing *R. opacus* culture grown to mid-exponential phase with 10%(v/v) glycerol solution. The stock was kept at -80°C and stored for no longer than 3 months. The pH of the shake flask cultures was maintained neutral by bolus addition of sodium hydroxide (NaOH) in response to changes in color of the pH indicator bromothymol blue, which was added to each flask at a concentration of 15 mg/L.

2.3 Batch and Fed-Batch Fermentations:

A New Brunswick BioFlo 110 fermentation system with 1-L capacity vessel was used for all batch experiments used for stoichiometry calculations and also for all fed-batch experiments. The defined media without the trace element solution, stock A solution, phosphate buffer or ammonium sulfate was autoclaved inside the vessel. Following the sterilization, the solutions were filter sterilized through a 0.2 μm membrane and added to the reactor through the injection port (septum). The pH was maintained at a setpoint of 6.9 using only sodium hydroxide (NaOH). The reaction was carried out at constant temperature of 30°C and air was

used to supply oxygen to the vessel at a rate of 1 vvm. The agitation was initialized at 300 rpm and it was increased by the controller as needed up to a maximum of 800 rpm to maintain dissolve oxygen levels above the setpoint of 20%.

2.4 Analysis Methods:

Total cell concentration is measured by lyophilizing the cells and weighing the dry cell mass. Residual dry cell weight is then estimated by subtracting the total fatty acid mass from the total cell dry weight. To determine the fatty acid content of the cells and the composition of the lipids, the lyophilized cells are subjected to methanolysis (Brandl et al 1988). The resulting fatty acid methyl esters (FAME) are subsequently analyzed by gas chromatography (GC-FAME) and the fatty acids are identified by comparison of their retention times with those of standard fatty acid methyl esters. The glucose concentration is measured in the supernatant, by high-performance liquid chromatography (HPLC) in an Agilent 1200 Series System using an Aminex Column and 5mM sulfuric acid as the mobile phase at a constant flowrate of 6mL/min and temperature of 55°C. The ammonium concentration is also measured in the supernatant, by enzymatic assay with L-glutamate dehydrogenase (Ammonia Assay Kit, Sigma-Aldrich Catalog No. AA0100). The assay calls for measurement of the decrease in absorbance at 340nm, due to the oxidation of NADPH, which in this case is proportional to ammonia concentration. The ammonia concentrations were then converted to ammonium sulfate concentrations and reported as the measurement for nitrogen content in the cultures.

CHAPTER 3
STOICHIOMETRY AND THE CARBON BALANCE

The stoichiometry of *Rhodococcus opacus* fermentations has been addressed, in particular the carbon balance. Using a gas analyzer we have been able to probe the off-gas from a bench scale 1L reactor, following the distribution of carbon in the system and closing the carbon balance of this process. In addition to stoichiometry, we have also taken aim at other important features of the process like establishing a correlation between cell dry weight and optical density. We have found that there is no apparent correlation between the optical density and the cell dry weight measured. As detailed in Figure 3, data from numerous experiments show that the linear correlation values vary radically from one *Rhodococcus* culture to the other. Therefore it is important to note that when working with *R. opacus* cultures, the only reliable numbers we can use to express cell concentration are cell dry weight numbers, whether they are determined by freeze drying or vacuum filter-drying. Both of these methods have been evaluated and found to yield comparable and reproducible results for cell dry weights.

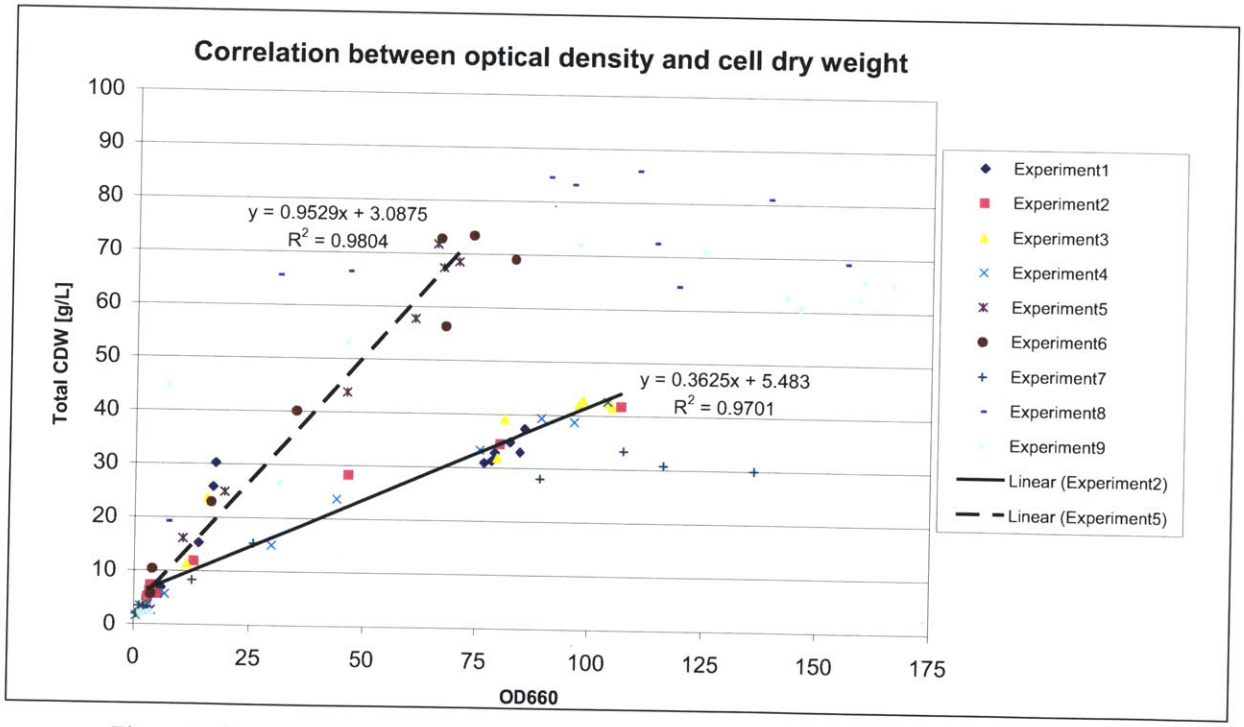

Figure 3: Comparison of nine individual experiments to establish the cell dry weight to optical density correlation for the organism *R. opacus* PD630. Two independent linear regressions are shown to highlight the dramatic variations in the correlation.

Using former batch data obtained at the Sinskey Lab using 0.5L scale Sixfors reactors, as well as elemental chemical composition analysis for lipid-containing and lipid-free *Rhodococcus opacus* we were able to establish two balanced stoichiometric equations for the fermentations on glucose. To generate these equations we calculated one parameter from experimental data: the yield of biomass on glucose and that was enough to fully determine the system and calculate all other stoichiometric coefficients.

"Fat" Rhodococcus:

$$C_6H_{12}O_6 + 1.86\ O_2 + 0.13\ NH_3 \rightarrow 0.63\ C_{5.1}H_{9.7}O_{1.6}N_{0.2} + 3.14\ H_2O + 2.79\ CO_2$$

"Lean" Rhodococcus:

$$C_6H_{12}O_6 + 2.58\ O_2 + 0.43\ NH_3 \rightarrow 0.72\ C_{4.2}H_8O_2N_{0.6} + 3.76\ H_2O + 2.98\ CO_2$$

Now if we focus only on the carbon balance for this *Rhodococcus* process, we can see that it can be expressed simply as moles of $C_{glucose}$ = moles of $C_{biomass}$ + moles of C_{CO_2}. Some earlier observations from flask experiments revealed the appearance of an unidentified peak in the HPLC chromatogram that was presumed to be an acid responsible for the dramatic pH drop in cultures without pH control. Because of the magnitude of the effect it was hypothesized that this acid could be accumulating in significant amounts and could be a significant carbon byproduct that would need to be included in our material balances. While we arranged for the necessary gas analyzer to measure off-gas carbon dioxide, we began addressing this issue by trying to identify the peak by comparison to standards of typical fermentation acids. In total, we screened for eight species: 3-hydroxybutyrate, butyric acid, gluconic acid, acetic acid, propionic acid, lactic acid, formic acid and phosphoric acid. Of these acids, only phosphoric acid closely approached the retention time of the target peak but it was not a perfect overlap. After further scrutiny, we were able to prove that the peak previously thought to be an acid was independent of pH and was present in cultures grown on either glucose or gluconate with pH ranges from 4 to neutral. This evidence refuted the acid byproduct hypothesis and we believe the peak can be attributed to phosphate species in the media/culture. The next level of analysis involved the use of a gas analyzer in batch fermentations at the 1L scale. The instrument used was an Agilent 3000 Micro GC and the calculations for cumulative carbon dioxide were performed using air as the reference gas. The

results presented in Figure 4 indicate that we can account for 91.7% of all carbon, with total biomass representing over 50% of all carbon and carbon dioxide being roughly 41%.

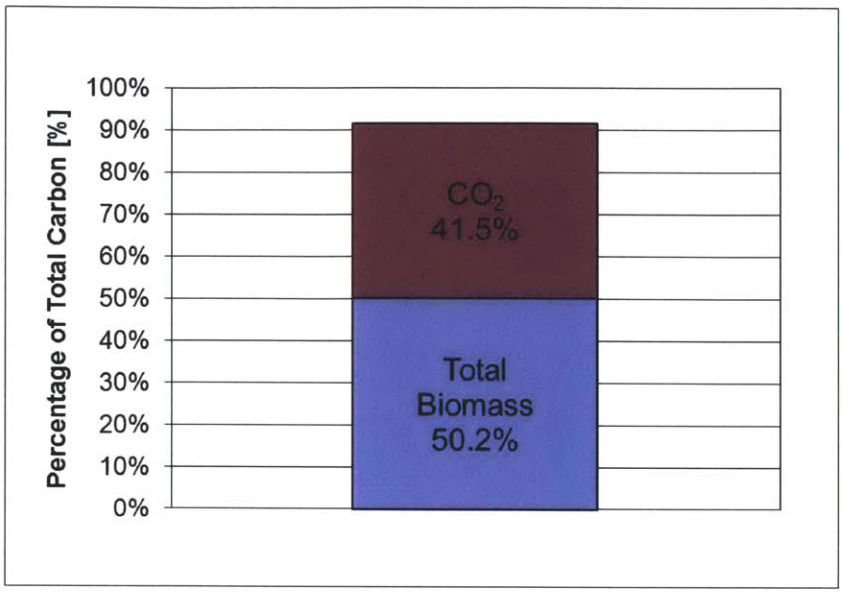

Figure 4: Carbon balance for a typical *Rhodococcus opacus* PD630 batch

It is also important to note that the distribution of carbon as function of time in the reactors remains relatively constant throughout the experiment. This supports the conclusion that the process does not generate significant amounts of organic acids or other carbon containing by-products. Note on Figure 5, that the standard deviations are rather large (maximum st.dev.= 9.5%) especially at initial and final time points. We believe that this large error comes mostly from instrument error and that it is possible for the carbon balance to be closed to greater than 95%. This hypothesis should be corroborated perhaps by using a mass spectrometer gas analyzer.

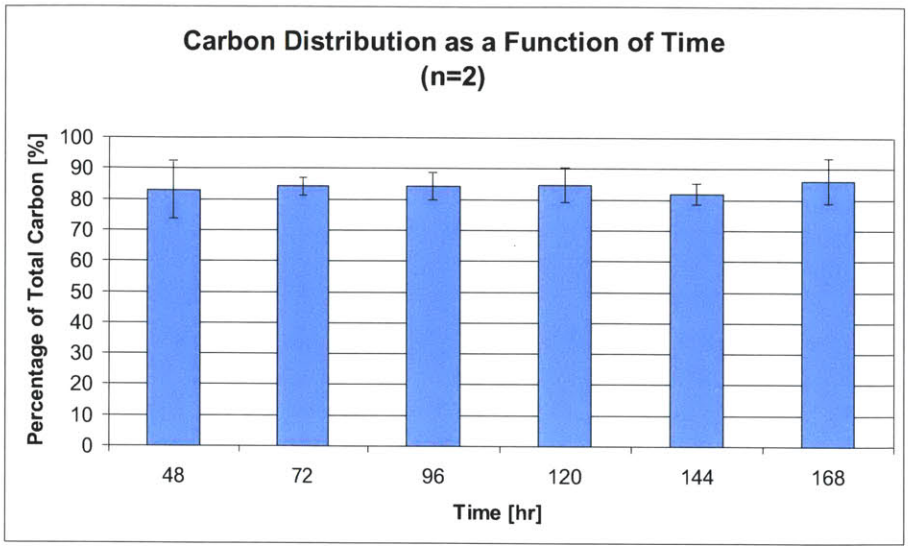

Figure 5: Carbon distribution as a function of time. Average values for Batches No. 8 and 9.

In addition to material balances, another important aspect of any biological or chemical process is the energy balance. To address the energy balance of this process we started with an enthalpy balance for microbial utilization of substrate as expressed in Equation 1; where the heat of combustion of the substrate is equal to the sum of the metabolic heat, the heat of combustion of biomass and the heat of combustion of the product.

$$S\left(\Delta H_S\right) = X\left(\Delta H_X + \frac{1}{Y_H}\right) + P\left(\Delta H_P\right) \qquad \text{Equation 1}$$

This expression is typically used to estimate the rate of heat evolution in batch fermentations and the ability to estimate this parameter is essential to proper reactor design as it is directly associated with assessment of heat removal requirements. Another way to think about energy is to define an efficiency term as the sum of all enthalpies of the products divided by the sum of enthalpies of the substrates. For these calculations we first need to obtain the heats of combustion of all species. Our substrate, glucose can be easily found in any textbook or reference manual. In the case of the TAG products, these heat values will vary depending on the chain length of the fatty acid species composing the TAG molecules. For our calculations we use an average between long-chain and medium-chain TAGs. The enthalpy of combustion of the "lean" *Rhodococcus* biomass was estimated from elemental composition analysis using available electron concepts (Patel and Erickson, 1981).

Species	Heat of combustion
TAG†	9 kcal/g
Glucose	3.8 kcal/g
Biomass*	5.4 kcal/g

†Average value, *Estimated

Table 1: Heats of combustion for substrate, biomass and triglyceride product

If we use the typical value of Y_H for glucose as 0.42 (Shuler and Kargi, 2002) in combination with our batch data we obtain an efficiency of **67.5% ± 7.5%**.

CHAPTER 4
FED-BATCH STRATEGY

In order to increase titers and productivity in biological processes two approaches can be used. The first involves metabolic engineering to genetically modify the host to outperform the wild type cells and the second follows classic reactor engineering concepts like changing a process configuration from batch to continuous operation. The latter approach was initially explored as we tried to develop a fed-batch strategy that could result in significant improvements in TAGs productivity using the wild type *Rhodococcus opacus* PD630 strain. The rationale for our strategy came from our interest in possibly making this process a continuous one, which would require both substrates, namely glucose and ammonium sulfate, to be fed simultaneously. While the condition of nitrogen limitation is still necessary for formation of product, numerous questions need to be answered in terms of the process limitations before the process can be converted to continuous operation. Given the requirement of nitrogen inhibition, identifying an optimal feed composition that will allow lipid accumulation will be critical. Other factors like feed timing and feed rate also need to be optimized. The experiments discussed in this chapter were designed to address these challenges and also provide some insight about the limits of this nitrogen starvation requirement. Figure 6 presents a summary with the results of three individual experiments with high, medium and low ammonium sulfate concentrations on the feed. The feed rate for all experiments was determined by equipment limitations (i.e., vessel volume) and remained constant at 7mL/hr. The feed was started only after the initial amount of ammonium sulfate substrate was exhausted in each reactor.

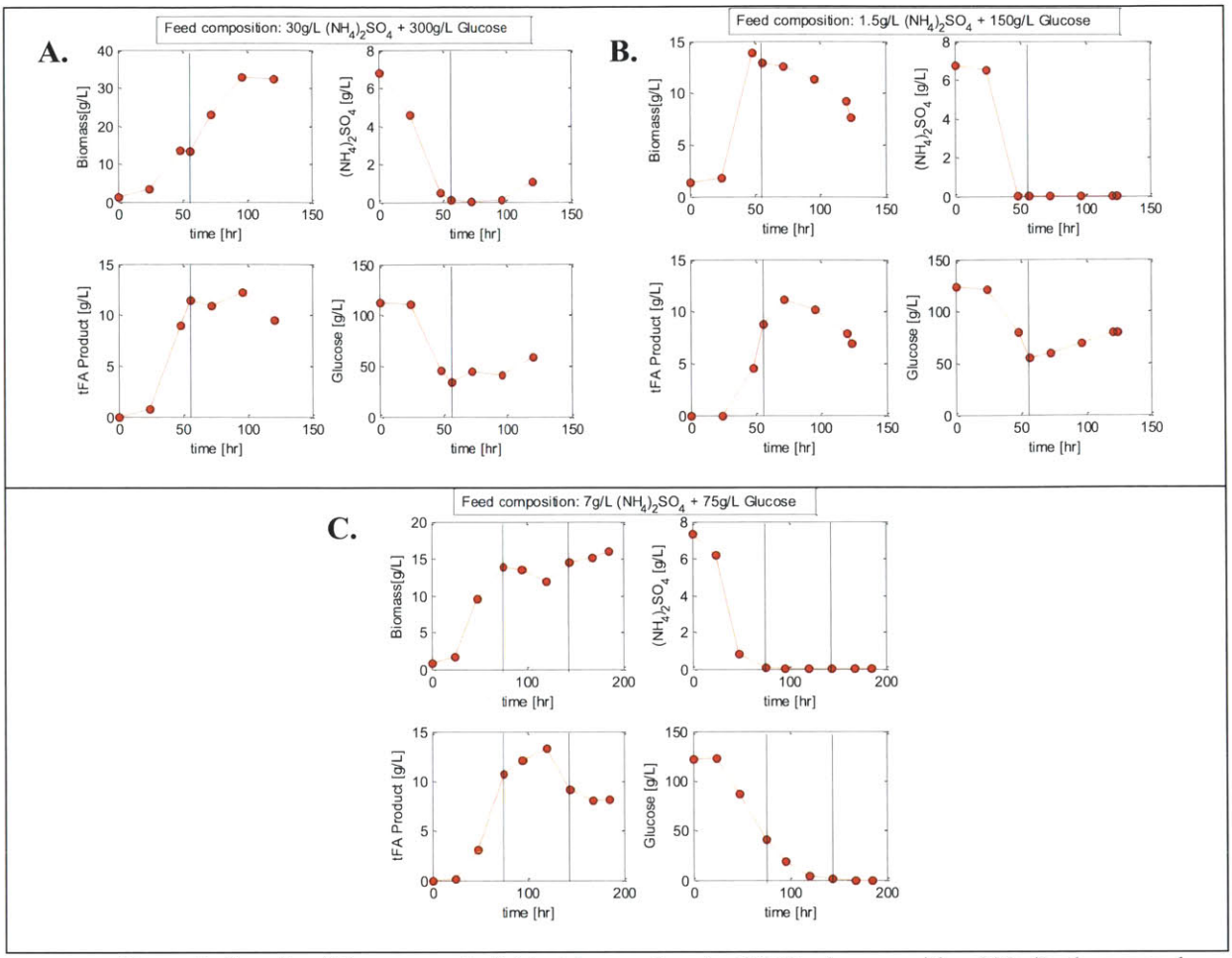

Figure 6: Results of three separate fed-batch experiments: (A) Feed composition 300g/L glucose and 30g/L ammonium sulfate at a rate of 7mL/hr. (B) Feed composition 150g/L glucose and 1.5g/L ammonium sulfate at a rate of 7mL/hr. (C) Feed composition 75g/L glucose and 7g/L ammonium sulfate at a rate of 7mL/hr. Dotted lines mark the feed start and/or stop time.

In the first two cases, we could attribute the drop in concentration of TAGs to dilution effects from the feed. However, the case presented in figure 6C, where the feed was stopped but the reaction was allowed to continue, provides evidence of TAG degradation. While TAG degradation is not dominant in batch cultures, it can happen when feeding is deficient and cells attempt to survive. In continuously fed cultures, lipid degradation will have a significant effect on product titers and productivity. Observations like these are critical for process design modifications and also provide physiological insights about potential targets for genetic modifications to *Rhodococcus opacus* strains. Both of these approaches can result in increasing yields and/or productivities for this process.

CHAPTER 5
SEED STUDIES

In most industrial settings seed culture preparations and reactor inoculations are part of Standard Operating Procedures (SOPs) documents. These are meant to provide uniformity to the process by minimizing human errors and ensuring reproducibility stays at its highest. Seed culture flasks are often started from seed bank vials, or cells that have been cryo-preserved uniformly and proceed from the same colony of cells. We conducted several seed studies before establishing a seed bank to be used to start all *Rhodococcus* cultures. Features like optimal inoculum age and inoculum size were evaluated to minimize lag phase in the cultures. Figure 7 shows the results of the inoculum age study in which a large "parent" flask culture is used at different stages of growth (i.e. early and late exponential phase) to inoculate several "daughter" flasks and analyze their corresponding growth behavior. The results indicate that cells have minimal lag if started from a culture in mid-exponential phase or later.

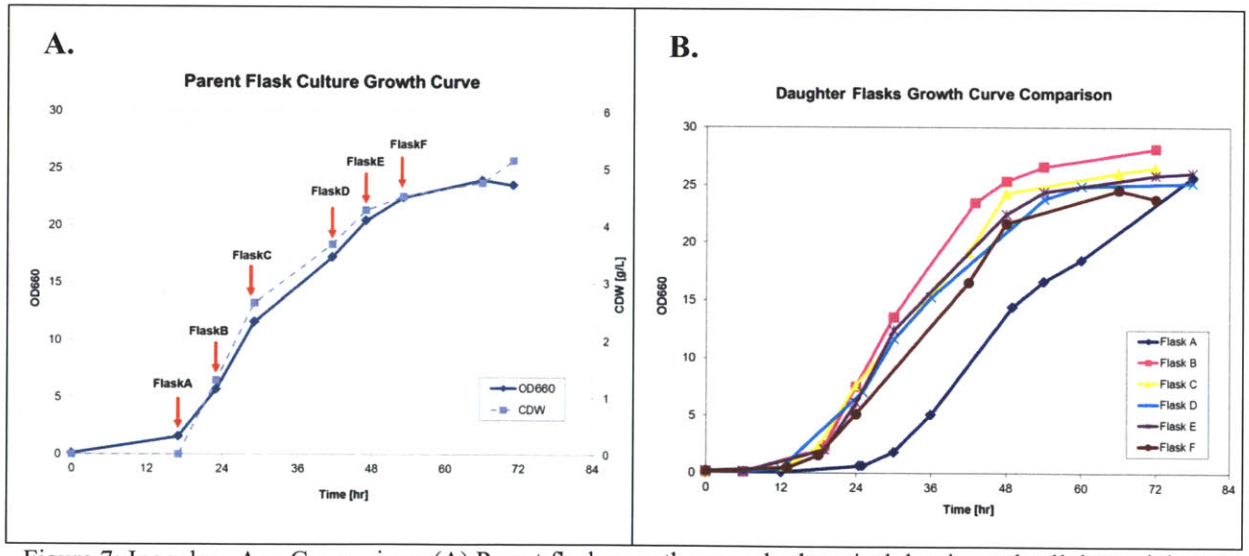

Figure 7: Inoculum Age Comparison. (A) Parent flask growth curve, both optical density and cell dry weight are used to measure cell concentration. Red arrows indicate times at which the culture was used to inoculate a new daughter culture. (B) Growth behavior of all daughter flasks compared.

Inoculum size refers to the volume of parent culture used to start a fresh daughter culture, and is typically expressed in volume percentages. For this analysis the objective is again to minimize the observed lag in the daughter cultures. As may be expected, higher inoculum

sizes will lead to less lag but a balance must be achieved keeping in mind that larger and larger volumes will be required to inoculate at reactor level during scale-up. For this purpose a range from 0.1% to 10% per volume of the same inoculum were used and the resulting growth curves are presented in Figure 8.

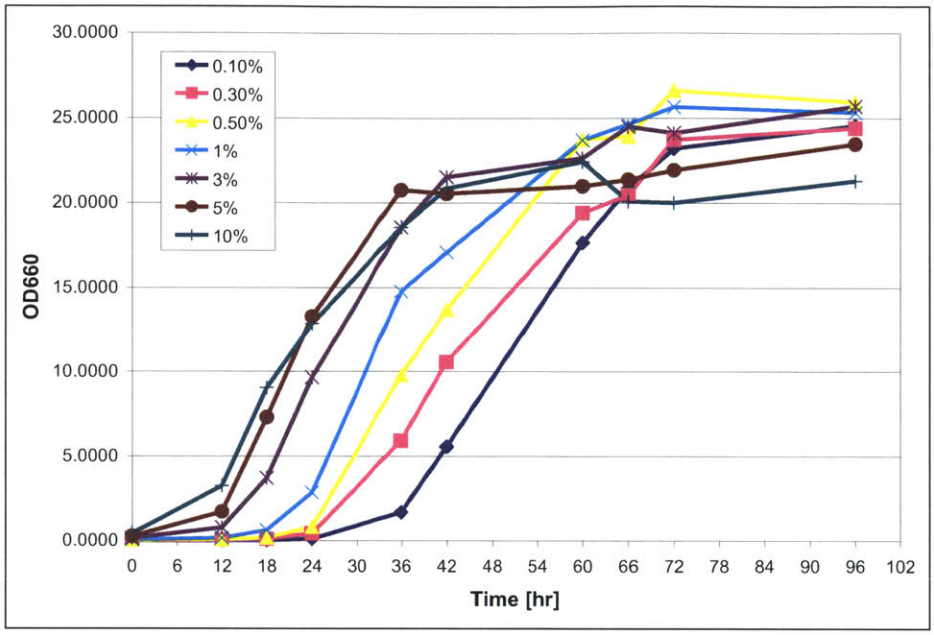

Figure 8: Effect of increasing inoculum sizes on growth behavior of *Rhodococcus opacus* PD630

Inoculums of at least 3% are recommended, as these bring the culture to a mid-exponential phase within 24 hours or less (see Figure 8), which is ideal for cultures to be used for reactor inoculation. Another aspect of seed preparation studied was the effect of the added cryoprotective additive on microbial physiology. Glycerol is among the most widely used cryoprotective additive for frozen storage of cells (Hubálek, 2003). The concentration of glycerol used varies slightly for different microorganisms, but in general a range of 20-30%vv stock solution and a ratio of 1:1 (glycerol stock : culture) is typically recommended for long term storage. When we compared the growth behavior of frozen cells suspended in 30% glycerol to the behavior of fresh cells, without any additive, a decrease in cell viability was observed (Figure 9).

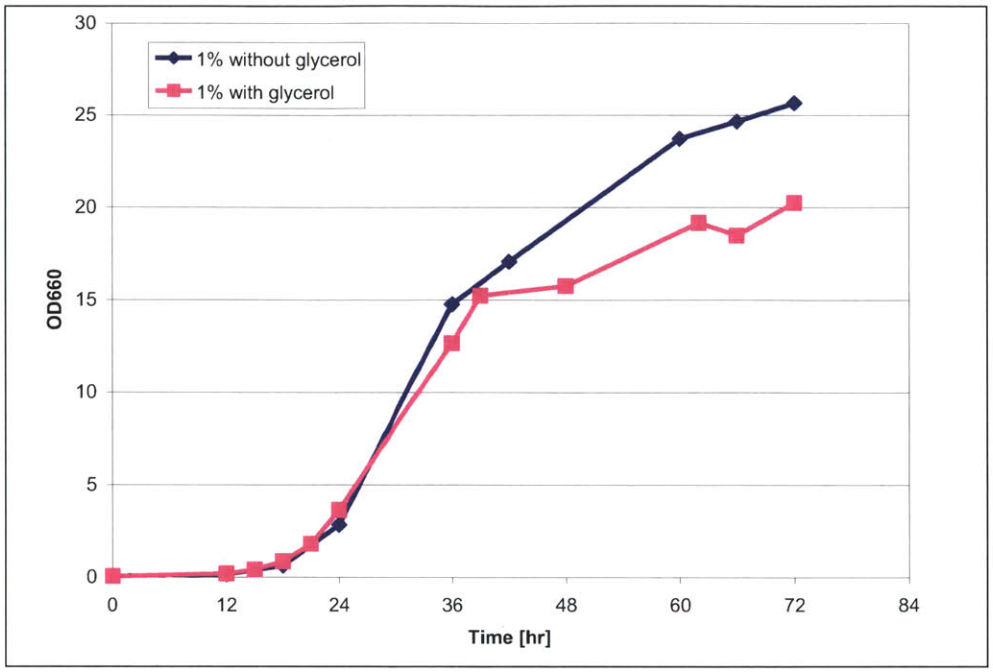

Figure 9: Effect of glycerol on *Rhodococcus opacus* PD630 growth performance

This effect on final cell titers can be explained by looking at substrate consumption of *Rhodococcus opacus* PD630 cell cultures, presented in Figure 10. We found evidence that suggest the cells are adjusting their metabolism to consume the glycerol present in the medium which is carried over from the frozen vials used for the culture flask inoculation. This metabolism shift has the potential to increase the lag phase of the reactor cultures for which these flasks are used as seed. Therefore it became a goal to reduce the carry-over of glycerol to the cultures by reducing the glycerol concentration of the stock solution used for cryopreservation.

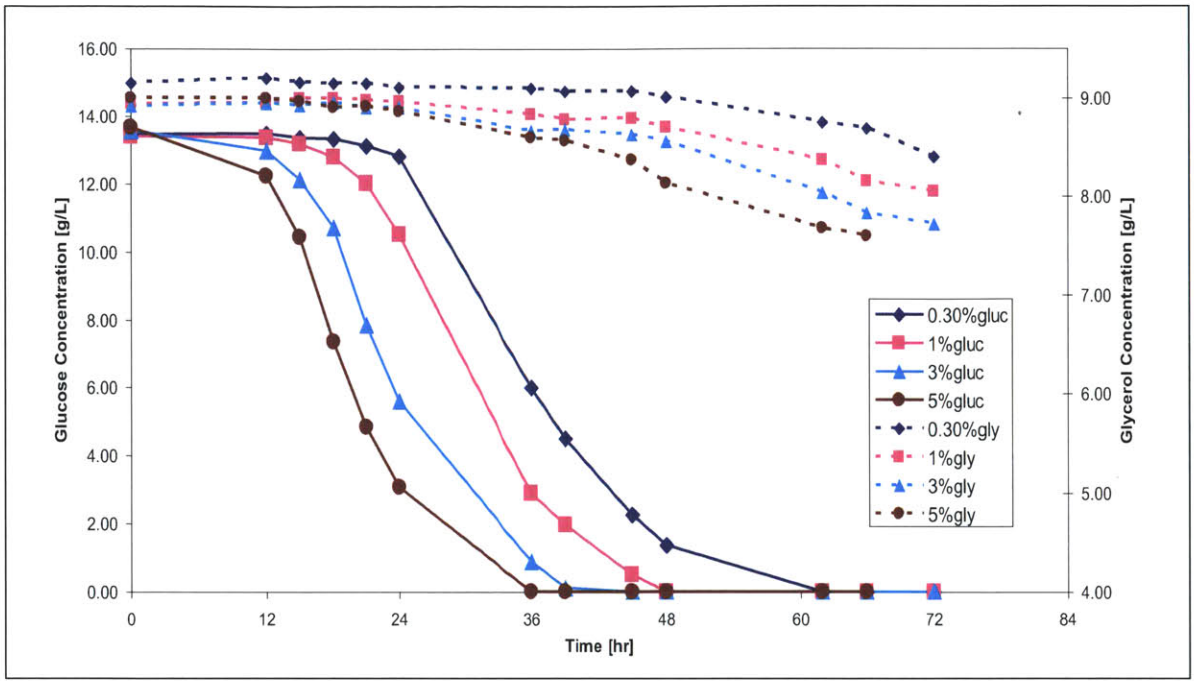

Figure 10: Consumption of glucose and glycerol by *Rhodococcus opacus* cultures started from frozen vials of cells cryopreserved using 30%vv glycerol.

The glycerol stock concentration was reduced from 30% to 10% and this new solution was evaluated in terms of its carryover effect and its cryopreservation capability. The concentration of the additive is a critical determinant of the cryopreservation efficacy over time. As a result, reducing the glycerol concentration will inevitably reduce the length of time that the seed bank frozen vials can be stored while maintaining optimal cell viability. The glycerol concentration reduction was effective in eliminating the deleterious effect on growth observed in Figure 9 yet the cell's viability after long term storage was drastically affected. The viability of the cells suspended in the low glycerol solution remained intact after 3 months of storage at -80°C, but started to deteriorate quickly after the 4th month of storage. Figure 11 shows that by 5 months the cells were lagging twice as long after 24 hours of cultivation at 30°C. Based on this observation a storage period no longer than 3 months is recommended for *Rhodococcus opacus* cells cryopreserved using 10% glycerol solution.

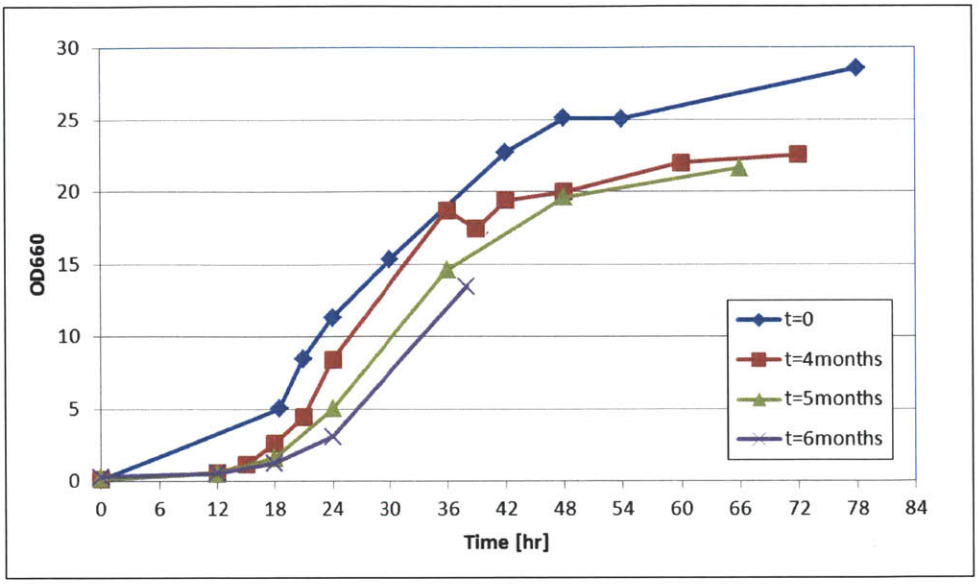

Figure 11: Performance of seed bank frozen cells evaluated over a period of six months.

CHAPTER 6
KINETIC MODEL DEVELOPMENT

When analyzing a system that is to be controlled or optimized engineers often use mathematical models. In general, models can be either descriptive or they can be used to try to estimate how an event or change will affect the process. This chapter details the development of a model that describes the production of TAGs in *Rhodococcus opacus* PD630.

The first species considered in the model is cell concentration, which is represented as residual biomass. This term, X, encompasses total biomass minus TAGs formed and is described by the following equation:

$$\frac{dX}{dt} = \mu X \qquad \text{Equation 2}$$

where, μ can be divided in two growth regimes associated with the substrate or substrates being used for growth during that time period. During growth on substrates, μ is described by a mixed substrate Monod relationship with respect to total ammonium sulfate and glucose concentrations:

$$\mu = \frac{\mu_{\max,S}(N * G)}{K + (N * G)} \qquad \text{Equation 3}$$

However, when glucose is completely exhausted and nitrogen is still present, the cells can adjust their metabolism to consume the TAGs they have accumulated and during this regime μ is in the form of the Monod relationship with respect to product (TAGs) concentration:

$$\mu = \frac{\mu_{\max,P}\ (N * P)}{K_P + (N * P)} \qquad \text{Equation 4}$$

The nitrogen concentration, N, is nitrogen measured as ammonium sulfate, $(NH_4)_2SO_4$, and is expected to be consumed as a function of biomass accumulation as shown in Equation 5:

$$\frac{dN}{dt} = -\frac{\mu X}{Y_{X/N}} \qquad \text{Equation 5}$$

where μ has the functionality described in Equations 3 and 4, and $Y_{X/N}$ is a yield coefficient of residual biomass on ammonium sulfate.

As we have observed TAGs being accumulated both during exponential growth and after cessation of growth, the product (TAGs) is described in this model with a mixed-growth associated product formation rate:

$$\frac{dP}{dt} = \alpha\mu X + \beta X \qquad \text{Equation 6}$$

where α is the growth associated product formation constant and β is the specific product formation rate, which is a function of glucose concentration, G, as defined by Equation 7:

$$\beta = \frac{\beta_{\max} G}{K_G + G} \qquad \text{Equation 7}$$

When glucose is exhausted and nitrogen is still available, the product accumulation stops and a product consumption phase begins in which the cells use the accumulated TAGs as a substrate. During this regime the product consumption rate is described by Equation 8:

$$\frac{dP}{dt} = -\frac{\mu X}{Y_{X/P}} \qquad \text{Equation 8}$$

where μ has the same functionality described in Equation 4 and $Y_{X/P}$ is a yield coefficient of residual biomass on product.

Finally, the glucose in the system is consumed both for cell growth and product formation as shown in Equation 9:

$$\frac{dG}{dt} = -\frac{(\alpha\mu X + \beta X)}{Y_{P/G}} - \frac{\mu X}{Y_{X/G}} \qquad \text{Equation 9}$$

where $Y_{P/G}$ is the yield of product on glucose and $Y_{X/G}$ is the yield of residual biomass on glucose.

The model equations described above produce a set of four simultaneous differential equations relating the residual biomass, glucose, ammonium sulfate and TAG concentrations. These differential equations are solved in MATLAB Version 7.11.0.584 (R2010b) using the "ode15s" solver (Appendix A.1). As part of the development of a complete kinetic model, comparisons between the proposed model and experimental data are used to obtain the

parameters for the model. The model described in Equations 2 through 9 contains a total of eleven constants shown in Table 2.

Parameter	*Value (Fitted by SSE minimization or measured)*	*95% Confidence Intervals (Estimated from nlinfit residuals and Jacobians)*
$Y_{X/N}$	1.8212 g biomass g $(NH_4)_2SO_4^{-1}$	0.9800 – 2.7698
$Y_{X/G}$	0.3377 g biomass g glucose^{-1}	0.1768 – 0.5017
$Y_{X/P}$	1.9382 g biomass g TAG^{-1}	0.1560 – 0.3100
$Y_{P/G}$	0.2320 g TAG g glucose^{-1}	1.1513 – 2.7082
$\mu_{max,S}$	0.185 hr^{-1}	N/A (measured)
$\mu_{max,P}$	0.0813 hr^{-1}	-0.1642 – 0.3178
β_{max}	0.2168 g TAG g biomass^{-1} hr^{-1}	-0.1057 – 0.5353
K	0.0023 g glucose g $(NH_4)_2SO_4$ L^{-2}	-0.0283 – 0.037
K_P	5.1013 g TAG L^{-1}	-23.5697 – 32.6466
K_G	24.7665 g glucose L^{-1}	-24.2060 – 76.6155
A	0.4075 g TAG g biomass^{-1}	0.1338 – 0.7503

Table 2: Model Parameters Summary (measured and regressed)

Of these constants, the maximum growth rate on substrates, $\mu_{max,S}$, is easily estimated from our experimental data as the slope of a semi-log plot of residual biomass concentration versus time. The remaining parameters were determined by minimization of the sum of squares for error (SSE) between the model's prediction and the observed growth, product formation and metabolite consumption profiles:

$$SSE = \sum_{i=1}^{n} \left(y_{\text{model}} - y_{\text{exp}} \right)^2 \qquad \text{Equation 10}$$

Data sets from five different experimental conditions, performed in triplicate at the shake flask level, were used for parameter determination (Figure 6). Initially, a Levenberg-Marquard algorithm was applied to perform least squares curve fitting using MATLAB's "nlinfit" function. However, this algorithm requires derivatives for all calculations, making it unsuitable for such a complex and highly correlated system of ODEs. Instead the parameter fit was done with the minimization algorithm known as Nelder-Mead simplex. This algorithm is included in MATLAB's "fminsearch" function and can be described as unconstrained non-

linear optimization based on heuristic rules. To find a global minimum the initial guesses for the parameters are varied over a range of possible parameter values and the solution accepted is that which results in the lowest value for the SSE. This exploration of parameter space was initiated with a set of parameter values fitted manually. Then 10% increases and decreases of each parameter were performed. With this iterative process the minimal SSE achieved was 91.229. To obtain the 95% confidence intervals presented in Table 2 the resulting parameters front the SSE minimization were used as initial guess inputs in a second MATLAB program using "nlinfit" (Appendix A.2), which calculates residuals and Jacobians that can subsequently be used by the MATLAB function "nlparci" to estimate the intervals.

From Figure 12 we can see that the model has problems predicting the behavior in flask C, which has 16g/L glucose and 8g/L ammonium sulfate. The model is over predicting the ammonium concentration and as a consequence the model also predicts a faster consumption of TAGs than observed experimentally. In all other cases, the model underestimates residual biomass. A way to address this problem is to evaluate the accuracy of the ammonium sulfate quantification method. The enzymatic method has been evaluated using ammonium sulfate standards however we have uncovered no significant inaccuracies in the quantification of these standards.

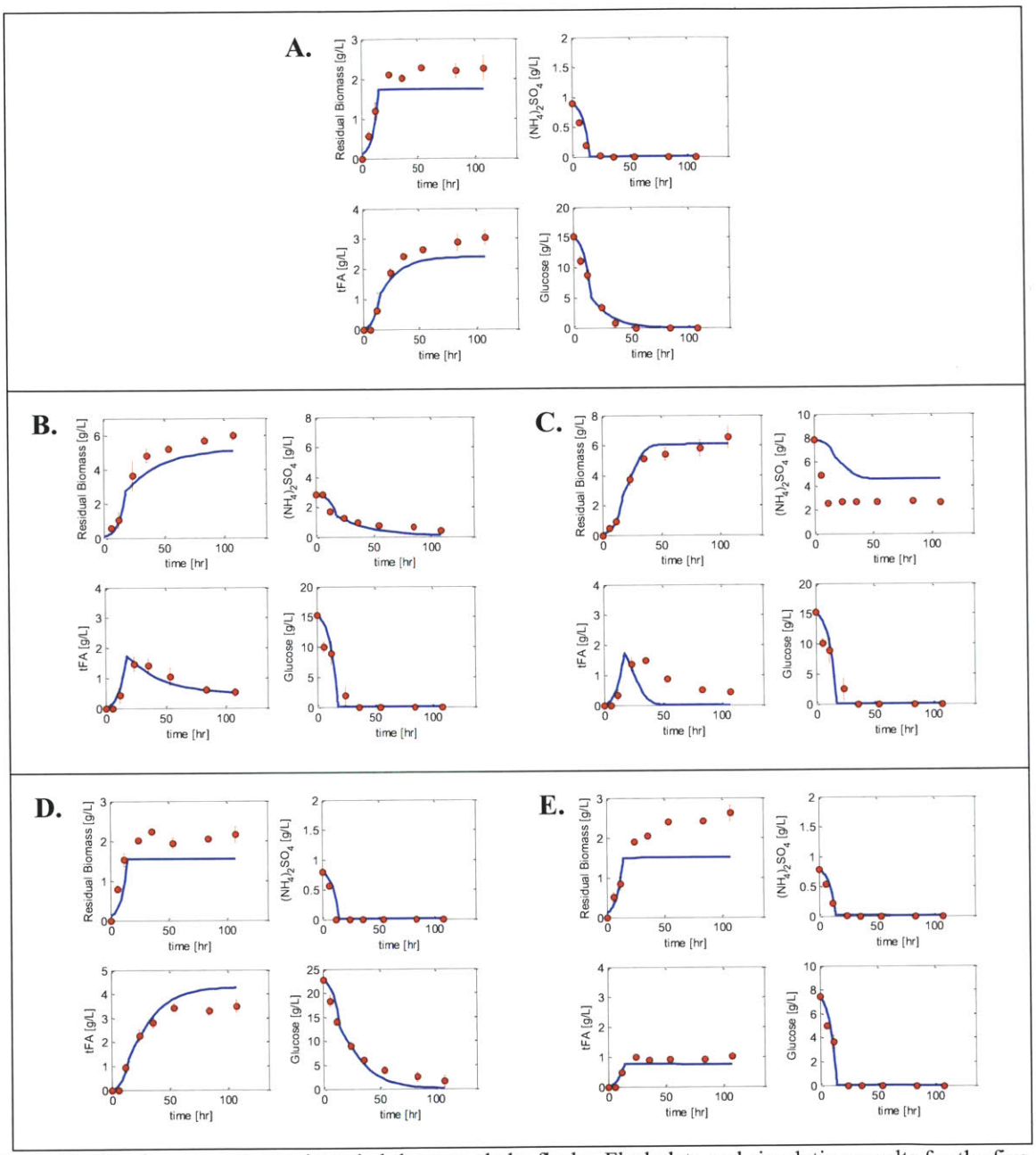

Figure 12: *Rhodococcus* process is scaled-down to shake flasks. Flask data and simulation results for the five conditions of ammonium sulfate and glucose concentration used in parameter estimation:

A: Basal flask medium containing 16g/L glucose and 1g/L ammonium sulfate.

B: Constant glucose, varying ammonia. The ammonium sulfate in this flask is increased to 4g/L, while keeping glucose at 16g/L.

C: Constant glucose, varying ammonia. The ammonium sulfate in this flask is further increased to 8g/L.

D: Constant ammonium sulfate, varying glucose. The glucose concentration is increased to 24g/L, while keeping ammonium sulfate at 1g/L.

E: Constant ammonium sulfate, varying glucose. Glucose is decreased to 8g/L.

CHAPTER 7
CONCLUSIONS AND RECOMMENDATIONS

Stoichiometric equations have been established for "fat" and "lean" *Rhodococcus opacus* PD630 according to their chemical elemental composition and experimental data. Off-gas reactor analysis data for carbon dioxide (CO_2) evolved demonstrates that the carbon balance is approximately 91% closed. Various fed-batch feeding strategies have been studied in order to maximize titers and productivity. However, the dual substrate feeding regime, with glucose and ammonium sulfate being supplied at different ratios, has not yet provided any evidence to support the hypothesis that fed-batch configuration can significantly increase the productivity of TAGs in *Rhodococcus opacus*. A systematic study of feed composition is recommended to either support or reject this hypothesis.

In another approach to increase productivity, seed cultures were optimized and a seed bank was created with low glycerol content. This minimizes both process time and lag time in the cultures, therefore increasing productivity. Finally, a kinetic model was developed to describe the production of triacylglycerols (TAGs) in *Rhodococcus opacus* PD630 under varying nutrient levels. The model adequately describes the observed growth and TAG accumulation, as well as its subsequent use as a substrate under experimental conditions of exhausted glucose and excess nitrogen. Nevertheless, we recommend the model be revisited and evaluated further, specifically with respect to ammonium sulfate concentrations. The current model advances our understanding of *Rhodococus opacus* PD630 process kinetics but further refinements to increase model robustness could ensure it becomes an essential tool in the development of applications for production of microbial oils with this organism.

REFERENCES

1. Alvarez, H. M. & Steinbuchel, A. Triacylglycerols in prokaryotic microorganisms. *Appl. Microbiol. Biotechnol.* **60**, 367-376 (2002).

2. Alvarez, H. M., Mayer, F., Fabritius, D. Formation of intracytoplasmic lipid inclusions by *Rhodococcus opacus* strain PD630. *Arch. Microbiol.* **165**, 377-386 (1996).

3. Asenjo, J. A. & Merchuk, J. C. Bioreactor system design. , 620 (1995).

4. Brandl, H., Gross, R.A., Lenz, R.W., Fuller, C. Pseudomonas oleovorans as a Source of Poly(β-Hydroxyalkanoates) for Potential Applications as Biodegradable Polyesters. *Appl.Environ.Microbiol.* **54**, 1977-1982 (1988).

5. Canakci, M. & Sanli, H. Biodiesel production from various feedstocks and their effects on the fuel properties. *J. Ind. Microbiol. Biotechnol.* **35**, 431-441 (2008).

6. Chartrain, M., Jackey, B., Taylor, C., Sandford, V., Gbewonyo, K., Lister, L., Dimichele, L., Hirsch, C., Heimbuch, B., Maxwell, C., Pascoe, D., Buckland, B., Greasham, R. Bioconversion of indene to cis (1S,2R) indandiol and trans (1R,2R) indandiol by Rhodococcus species. *J. Ferment. Bioeng.* **86**, 550–558 (1998).

7. Cramer, A. C., Vlassides, S., Block, D. E. Kinetic model for nitrogen-limited wine fermentations *Biotechnol. Bioeng.* **77**, 49-60 (2002).

8. Durrett, T. P., Benning, C., Ohlrogge, J. Plant triacylglycerols as feedstocks for the production of biofuels. *Plant J.* **54**, 593-607 (2008).

9. Elbahloul, Y. & Steinbuchel, A. Pilot-scale production of fatty acid ethyl esters by an engineered *Escherichia coli* strain harboring the p(Microfiesel) plasmid. *Appl. Environ. Microbiol.* **76**, 4560-4565 (2010).

10. Hernández, M. A., Mohn, W.W., Martínez, E., Rost, E., Alvarez, A.F., Alvarez, H.M. Biosynthesis of storage compounds by *Rhodococcus jostii* RHAI global identification of genes involved in their metabolism. *BMC Genomics*, **9**:600 (2008).

11. Hubálek, Z. Protectants used in the cryopreservation of microorganisms. *Cryobiology*. **46**, 205-229 (2003).

12. Kurosawa, K., Boccazzi, P., de Almeida, N.M., Sinskey, A.J. High-cell density batch fermentation of *Rhodococcus opacus* PD630 using a high glucose concentration for triacylglycerol production. *J. Biotech*. **147**, 212-218 (2010).

13. Li, Q., Du, W., Liu, D. Perspectives of microbial oils for biodiesel production. *Appl. Microbiol. Biotechnol.* **80**, 749-756 (2008).

14. Murphy, D. J. & Vance, J. Mechanisms of lipid-body formation *Trends Biochem. Sci.* **24**, 109-115 (1999).

15. Patel, S. A. & Erickson, L. E. Estimation of heats of combustion of biomass from elemental analysis using available electron concepts. *Biotechnol. Bioeng.* **23**, 2051-2067 (1981).

16. Shuler, M. L. & Kargi, F. *Bioprocess Engineering: Basic Concepts* 2nd Ed. (Prentice Hall, Upper Saddle River, NJ, 2002).

17. Vasudevan, P. T. & Briggs, M. Biodiesel production--current state of the art and challenges. *J. Ind. Microbiol. Biotechnol.* **35**, 421-430 (2008).

18. Waltermann, M., Hinz, A., Robenek, H., Troyer, D., Reichelt, R., Malkus, U., Galla, H.J., Kalscheuer, R., Stöveken, T., von Landenberg, P., Steinbuchel, A. Mechanism of lipid-

body formation in prokaryotes: how bacteria fatten up. *Mol. Microbiol.* **55**, 750-763 (2005).

19. Waltermann, M., Luftmann, H., Baumeister, D., Kalscheuer, R., Steinbuchel, A. Rhodococcus opacus strain PD630 as a new source of high-value single-cell oil? Isolation and characterization of triacylglycerols and other storage lipids. *Microbiology* **146**, 1143-1149 (2000).

20. Waltermann, M. & Steinbuchel, A. Neutral lipid bodies in prokaryotes: recent insights into structure, formation, and relationship to eukaryotic lipid depots. *J. Bacteriol.* **187**, 3607-3619 (2005).

APPENDIX

A.1: Matlab code for *Rhodococcus opacus* model parameter estimation

```
function [flag]=parameterminallflasknolag();
%This function finds the best parameters to fit the batch
R.opacus model
%simulation to the data by minimizing the sum of squared
errors (sse).
clc; close all;
flag=0;
%Read experimental data from excel file
EXP1=xlsread('flaskdataK.xls',1,'C3:K10');
EXP2=xlsread('flaskdataK.xls',1,'C12:K19');
EXP3=xlsread('flaskdataK.xls',1,'C21:K28');
EXP4=xlsread('flaskdataK.xls',1,'C30:K37');
EXP5=xlsread('flaskdataK.xls',1,'C39:K46');
t1=EXP1(:,1);
X1=EXP1(:,2);
X1error=EXP1(:,3);
N1=EXP1(:,4);
N1error=EXP1(:,5);
P1=EXP1(:,6);
P1error=EXP1(:,7);
G1=EXP1(:,8);
G1error=EXP1(:,9);
Y1=[X1;P1;G1];
t2=EXP2(:,1);
X2=EXP2(:,2);
X2error=EXP2(:,3);
N2=EXP2(:,4);
N2error=EXP2(:,5);
P2=EXP2(:,6);
P2error=EXP2(:,7);
G2=EXP2(:,8);
G2error=EXP2(:,9);
Y2=[X2;P2;G2];
t3=EXP3(:,1);
X3=EXP3(:,2);
X3error=EXP3(:,3);
N3=EXP3(:,4);
N3error=EXP3(:,5);
P3=EXP3(:,6);
P3error=EXP3(:,7);
G3=EXP3(:,8);
G3error=EXP3(:,9);
Y3=[X3;P3;G3];
```

```
t4=EXP4(:,1);
X4=EXP4(:,2);
X4error=EXP4(:,3);
N4=EXP4(:,4);
N4error=EXP4(:,5);
P4=EXP4(:,6);
P4error=EXP4(:,7);
G4=EXP4(:,8);
G4error=EXP4(:,9);
Y4=[X4;P4;G4];
t5=EXP5(:,1);
X5=EXP5(:,2);
X5error=EXP5(:,3);
N5=EXP5(:,4);
N5error=EXP5(:,5);
P5=EXP5(:,6);
P5error=EXP5(:,7);
G5=EXP5(:,8);
G5error=EXP5(:,9);
Y5=[X5;P5;G5];

Data=[Y1;Y2;Y3;Y4;Y5];
knobs=[[t1;N1(1);G1(1)];[t2;N2(1);G2(1)];[t3;N3(1);G3(1)];[t4;
N4(1);G4(1)];[t5;N5(1);G5(1)]]; %Dummy variable carrying
inputs needed for other functions
Starting=[1.8212 0.3377 0.2320 0.2168 24.7665 0.0029 0.4075
1.9382 0.0813 5.1013]; %Initial guesses for parameters
options=optimset('Display','iter','MaxIter',1,'MaxFunEvals',40
00);
Estimates=fminsearch(@myfit,Starting,options,knobs,Data);
%Display fitted parameters result
{'Parameter','Value';
'Yxn',[Estimates(1)];
'Yxs',[Estimates(2)];
'Yps',[Estimates(3)];
'bmax',[Estimates(4)];
'Kg',[Estimates(5)];
'K',[Estimates(6)];
'a',[Estimates(7)];
'Yxp',[Estimates(8)];
'umaxP',[Estimates(9)];
'Kp',[Estimates(10)]}

%To check the fit plot data and simulation with estimated
parameters
%Call ode using estimates as inputs
for nexp=1:5
```

```
    if nexp==1
        N0=knobs(9);
        G0=knobs(10);
        Ci0=[0.1 N0 0 G0];
        [ti,Ci]=ode15s(@(ti,Ci) batch(ti,Ci,Estimates),[0
108],Ci0);
        figure(nexp)
        subplot(2,2,1),plot(ti,Ci(:,1),'LineWidth',2); hold on

subplot(2,2,1),errorbar(t1,X1,X1error,'ro','MarkerEdgeColor','
k','MarkerFaceColor','r','MarkerSize',5);
        subplot(2,2,1),xlabel('time [hr]');
        subplot(2,2,1),ylabel('Residual Biomass [g/L]');
        subplot(2,2,1),axis([0 136 0 3]);
        subplot(2,2,2),plot(ti,Ci(:,2),'LineWidth',2); hold on

subplot(2,2,2),plot(t1,N1,'ro','MarkerEdgeColor','k','MarkerFa
ceColor','r','MarkerSize',5);

%subplot(2,2,2),errorbar(texp,N,Nerror,'ro','MarkerEdgeColor',
'k','MarkerFaceColor','r','MarkerSize',5);
        subplot(2,2,2),xlabel('time [hr]');
        subplot(2,2,2),ylabel('(NH_4)_2SO_4 [g/L]');
        subplot(2,2,2),axis([0 136 0 2]);
        subplot(2,2,3),plot(ti,Ci(:,3),'LineWidth',2); hold on

subplot(2,2,3),errorbar(t1,P1,P1error,'ro','MarkerEdgeColor','
k','MarkerFaceColor','r','MarkerSize',5);
        subplot(2,2,3),xlabel('time [hr]');
        subplot(2,2,3),ylabel('tFA [g/L]');
        subplot(2,2,3),axis([0 136 0 4]);
        subplot(2,2,4),plot(ti,Ci(:,4),'LineWidth',2); hold on

subplot(2,2,4),errorbar(t1,G1,G1error,'ro','MarkerEdgeColor','
k','MarkerFaceColor','r','MarkerSize',5);
        subplot(2,2,4),xlabel('time [hr]');
        subplot(2,2,4),ylabel('Glucose [g/L]');
        subplot(2,2,4),axis([0 136 0 20]);
    else if nexp==2
            N0=knobs(19);
            G0=knobs(20);
            Ci0=[0.1 N0 0 G0];
            [ti,Ci]=ode15s(@(ti,Ci) batch(ti,Ci,Estimates),[0
108],Ci0);
            figure(nexp)
            subplot(2,2,1),plot(ti,Ci(:,1),'LineWidth',2);
hold on
```

```
subplot(2,2,1),errorbar(t2,X2,X2error,'ro','MarkerEdgeColor','
k','MarkerFaceColor','r','MarkerSize',5);
            subplot(2,2,1),xlabel('time [hr]');
            subplot(2,2,1),ylabel('Residual Biomass [g/L]');
            subplot(2,2,1),axis([0 136 0 7]);
            subplot(2,2,2),plot(ti,Ci(:,2),'LineWidth',2);
hold on

subplot(2,2,2),plot(t2,N2,'ro','MarkerEdgeColor','k','MarkerFa
ceColor','r','MarkerSize',5);

%subplot(2,2,2),errorbar(texp,N,Nerror,'ro','MarkerEdgeColor',
'k','MarkerFaceColor','r','MarkerSize',5);
            subplot(2,2,2),xlabel('time [hr]');
            subplot(2,2,2),ylabel('(NH_4)_2SO_4 [g/L]');
            subplot(2,2,2),axis([0 136 0 8]);
            subplot(2,2,3),plot(ti,Ci(:,3),'LineWidth',2);
hold on

subplot(2,2,3),errorbar(t2,P2,P2error,'ro','MarkerEdgeColor','
k','MarkerFaceColor','r','MarkerSize',5);
            subplot(2,2,3),xlabel('time [hr]');
            subplot(2,2,3),ylabel('tFA [g/L]');
            subplot(2,2,3),axis([0 136 0 4]);
            subplot(2,2,4),plot(ti,Ci(:,4),'LineWidth',2);
hold on

subplot(2,2,4),errorbar(t2,G2,G2error,'ro','MarkerEdgeColor','
k','MarkerFaceColor','r','MarkerSize',5);
            subplot(2,2,4),xlabel('time [hr]');
            subplot(2,2,4),ylabel('Glucose [g/L]');
            subplot(2,2,4),axis([0 136 0 20]);
        else if nexp==3
                N0=knobs(29);
                G0=knobs(30);
                Ci0=[0.1 N0 0 G0];
                [ti,Ci]=ode15s(@(ti,Ci)
batch(ti,Ci,Estimates),[0 108],Ci0);
                figure(nexp)
                subplot(2,2,1),plot(ti,Ci(:,1),'LineWidth',2);
hold on

subplot(2,2,1),errorbar(t3,X3,X3error,'ro','MarkerEdgeColor','
k','MarkerFaceColor','r','MarkerSize',5);
                subplot(2,2,1),xlabel('time [hr]');
```

```
                subplot(2,2,1),ylabel('Residual Biomass
[g/L]');
                subplot(2,2,1),axis([0 136 0 8]);
                subplot(2,2,2),plot(ti,Ci(:,2),'LineWidth',2);
hold on

subplot(2,2,2),plot(t3,N3,'ro','MarkerEdgeColor','k','MarkerFa
ceColor','r','MarkerSize',5);

%subplot(2,2,2),errorbar(texp,N,Nerror,'ro','MarkerEdgeColor',
'k','MarkerFaceColor','r','MarkerSize',5);
                subplot(2,2,2),xlabel('time [hr]');
                subplot(2,2,2),ylabel('(NH_4)_2SO_4 [g/L]');
                subplot(2,2,2),axis([0 136 0 10]);
                subplot(2,2,3),plot(ti,Ci(:,3),'LineWidth',2);
hold on

subplot(2,2,3),errorbar(t3,P3,P3error,'ro','MarkerEdgeColor','
k','MarkerFaceColor','r','MarkerSize',5);
                subplot(2,2,3),xlabel('time [hr]');
                subplot(2,2,3),ylabel('tFA [g/L]');
                subplot(2,2,3),axis([0 136 0 4]);
                subplot(2,2,4),plot(ti,Ci(:,4),'LineWidth',2);
hold on

subplot(2,2,4),errorbar(t3,G3,G3error,'ro','MarkerEdgeColor','
k','MarkerFaceColor','r','MarkerSize',5);
                subplot(2,2,4),xlabel('time [hr]');
                subplot(2,2,4),ylabel('Glucose [g/L]');
                subplot(2,2,4),axis([0 136 0 20]);
            else if nexp==4
                    N0=knobs(39);
                    G0=knobs(40);
                    Ci0=[0.1 N0 0 G0];
                    [ti,Ci]=ode15s(@(ti,Ci)
batch(ti,Ci,Estimates),[0 108],Ci0);
                    figure(nexp)

subplot(2,2,1),plot(ti,Ci(:,1),'LineWidth',2); hold on

subplot(2,2,1),errorbar(t4,X4,X4error,'ro','MarkerEdgeColor','
k','MarkerFaceColor','r','MarkerSize',5);
                    subplot(2,2,1),xlabel('time [hr]');
                    subplot(2,2,1),ylabel('Residual Biomass
[g/L]');
                    subplot(2,2,1),axis([0 136 0 3]);
```

```
subplot(2,2,2),plot(ti,Ci(:,2),'LineWidth',2); hold on

subplot(2,2,2),plot(t4,N4,'ro','MarkerEdgeColor','k','MarkerFa
ceColor','r','MarkerSize',5);

%subplot(2,2,2),errorbar(texp,N,Nerror,'ro','MarkerEdgeColor',
'k','MarkerFaceColor','r','MarkerSize',5);
                    subplot(2,2,2),xlabel('time [hr]');
                    subplot(2,2,2),ylabel('(NH_4)_2SO_4
[g/L]');
                    subplot(2,2,2),axis([0 136 0 2]);

subplot(2,2,3),plot(ti,Ci(:,3),'LineWidth',2); hold on

subplot(2,2,3),errorbar(t4,P4,P4error,'ro','MarkerEdgeColor','
k','MarkerFaceColor','r','MarkerSize',5);
                    subplot(2,2,3),xlabel('time [hr]');
                    subplot(2,2,3),ylabel('tFA [g/L]');
                    subplot(2,2,3),axis([0 136 0 5]);

subplot(2,2,4),plot(ti,Ci(:,4),'LineWidth',2); hold on

subplot(2,2,4),errorbar(t4,G4,G4error,'ro','MarkerEdgeColor','
k','MarkerFaceColor','r','MarkerSize',5);
                    subplot(2,2,4),xlabel('time [hr]');
                    subplot(2,2,4),ylabel('Glucose [g/L]');
                    subplot(2,2,4),axis([0 136 0 25]);
                else if nexp==5
                        N0=knobs(49);
                        G0=knobs(50);
                        Ci0=[0.1 N0 0 G0];
                        [ti,Ci]=ode15s(@(ti,Ci)
batch(ti,Ci,Estimates),[0 108],Ci0);
                        figure(nexp)

subplot(2,2,1),plot(ti,Ci(:,1),'LineWidth',2); hold on

subplot(2,2,1),errorbar(t5,X5,X5error,'ro','MarkerEdgeColor','
k','MarkerFaceColor','r','MarkerSize',5);
                        subplot(2,2,1),xlabel('time [hr]');
                        subplot(2,2,1),ylabel('Residual
Biomass [g/L]');
                        subplot(2,2,1),axis([0 136 0 3]);

subplot(2,2,2),plot(ti,Ci(:,2),'LineWidth',2); hold on
```

```
subplot(2,2,2),plot(t5,N5,'ro','MarkerEdgeColor','k','MarkerFa
ceColor','r','MarkerSize',5);

%subplot(2,2,2),errorbar(texp,N,Nerror,'ro','MarkerEdgeColor',
'k','MarkerFaceColor','r','MarkerSize',5);
                        subplot(2,2,2),xlabel('time [hr]');
                        subplot(2,2,2),ylabel('(NH_4)_2SO_4
[g/L]');
                        subplot(2,2,2),axis([0 136 0 2]);

subplot(2,2,3),plot(ti,Ci(:,3),'LineWidth',2); hold on

subplot(2,2,3),errorbar(t5,P5,P5error,'ro','MarkerEdgeColor','
k','MarkerFaceColor','r','MarkerSize',5);
                        subplot(2,2,3),xlabel('time [hr]');
                        subplot(2,2,3),ylabel('tFA [g/L]');
                        subplot(2,2,3),axis([0 136 0 4]);

subplot(2,2,4),plot(ti,Ci(:,4),'LineWidth',2); hold on

subplot(2,2,4),errorbar(t5,G5,G5error,'ro','MarkerEdgeColor','
k','MarkerFaceColor','r','MarkerSize',5);
                        subplot(2,2,4),xlabel('time [hr]');
                        subplot(2,2,4),ylabel('Glucose
[g/L]');
                        subplot(2,2,4),axis([0 136 0 10]);
                end
            end
        end
     end
    end
end
flag=1;
return;
%================================================================
=============
function sse=myfit(params,Input,Actual_Output)
Ymodel=0;
for nexp=1:5
    if nexp==1
        timevec=Input(1:8);
        N0=Input(9);
        G0=Input(10);
    else if nexp==2
            timevec=Input(11:18);
            N0=Input(19);
```

```
            G0=Input(20);
        else if nexp==3
                timevec=Input(21:28);
                N0=Input(29);
                G0=Input(30);
            else if nexp==4
                    timevec=Input(31:38);
                    N0=Input(39);
                    G0=Input(40);
                else if nexp==5
                        timevec=Input(41:48);
                        N0=Input(49);
                        G0=Input(50);
                end
            end
        end
    end
 end
%Initial concentrations (X, N, P, G)
Y0=[0.1 N0 0 G0];
%Call ode solver
options=odeset('RelTol',1e-1,'AbsTol',1e-3);
[t,that]=ode15s(@(t,Data)
simul(t,Data,params),timevec,Y0,options);
Ymodel=[Ymodel;[that(:,1);that(:,3);that(:,4)]];
end
Fitted=Ymodel(2:length(Ymodel));
%Calculate sum of squared errors
Error_Vector=Fitted - Actual_Output;
sse=sum(Error_Vector.^2);

return;
%================================================================
==========
function f = simul(t,Data,params)

%Define parameters for rate constant calculations
Yxn=params(1);
Yxs=params(2);
Yps=params(3);
umaxS=0.185;
bmax=params(4);
Kg=params(5);
K=params(6);
a=params(7);
Yxp=params(8);
umaxP=params(9);
```

```
Kp=params(10);

%Define species concentrations
X = Data(1);
N = Data(2);
P = Data(3);
G = Data(4);

%Define u and b
if (G>0.0001 && N>0.0001)
    u=(umaxS*N*G)/(K+(N*G));
elseif (G<0.0001 && N>0.0001)
    u=(umaxP*P*N)/(Kp+(P*N));
else
    u=0;
end
b=(bmax*G)/(Kg+G);

%X RESIDUAL BIOMASS TERM
rX = u*X;
%N NITROGEN CONSUMPTION TERM
if N>0.0001
    rN = -(u*X)/Yxn;
else
   rN=0;
   N=0;
end
%P PRODUCT FORMATION AND CONSUMPTION TERM
if (G>0.0001)
    rP = (a*u*X)+(b*X);
else
    rP=-u*X/Yxp;
end
%G GLUCOSE CONSUMPTION TERM
if G>0.0001
    rG = ((-a*u*X)/Yps)-((b*X)/Yps)-((u*X)/Yxs);
else
   rG=0;
   G=0;
end

f=zeros(4,1);
f(1)=rX;
f(2)=rN;
f(3)=rP;
f(4)=rG;
```

```
return;
%================================================================
=============
function y = batch(ti,Ci,Fitted_Params)
%Define parameters
Yxn=Fitted_Params(1);
Yxs=Fitted_Params(2);
Yps=Fitted_Params(3);
umax=0.185;
bmax=Fitted_Params(4);
Kg=Fitted_Params(5);
K=Fitted_Params(6);
a=Fitted_Params(7);
Yxp=Fitted_Params(8);
g=Fitted_Params(9);
Kp=Fitted_Params(10);

%Define species concentrations
X = Ci(1);
N = Ci(2);
P = Ci(3);
G = Ci(4);

%Define u and b
if (G>0.0001 && N>0.0001)
    u=(umax*N*G)/(K+(N*G));
elseif (G<0.0001 && N>0.0001)
    u=(g*P*N)/(Kp+(P*N));
else
    u=0;
end
b=(bmax*G)/(Kg+G);

%X RESIDUAL BIOMASS TERM
rX = u*X;
%N NITROGEN CONSUMPTION TERM
if N>0.0001
    rN = -(u*X)/Yxn;
else
   rN=0;
   N=0;
end
%P PRODUCT FORMATION AND CONSUMPTION TERM
if (G>0.0001)
    rP = (a*u*X)+(b*X);
else
    rP=-u*X/Yxp;
```

```
end
%G GLUCOSE CONSUMPTION TERM
if G>0.0001
    rG = ((-a*u*X)/Yps)-((b*X)/Yps)-((u*X)/Yxs);
else
   rG=0;
   G=0;
end

y=zeros(4,1);
y(1)=rX;
y(2)=rN;
y(3)=rP;
y(4)=rG;

return;
```

A.2: Matlab code used for *Rhodococcus opacus* model confidence interval estimations

```
function flagmain = paramfitflask()
%This function estimates the coefficients of our Rhodococcus
batch model using
%non-linear regression.
flagmain=0;
clear all; close all; clc;

%Read experimental data from excel file
%Read experimental data from excel file
EXP1=xlsread('flaskdataK.xls',1,'C3:K10');
EXP2=xlsread('flaskdataK.xls',1,'C12:K19');
EXP3=xlsread('flaskdataK.xls',1,'C21:K28');
EXP4=xlsread('flaskdataK.xls',1,'C30:K37');
EXP5=xlsread('flaskdataK.xls',1,'C39:K46');
t1=EXP1(:,1);
X1=EXP1(:,2);
X1error=EXP1(:,3);
N1=EXP1(:,4);
N1error=EXP1(:,5);
P1=EXP1(:,6);
P1error=EXP1(:,7);
G1=EXP1(:,8);
G1error=EXP1(:,9);
Y1=[X1;P1;G1];
t2=EXP2(:,1);
X2=EXP2(:,2);
X2error=EXP2(:,3);
N2=EXP2(:,4);
N2error=EXP2(:,5);
P2=EXP2(:,6);
P2error=EXP2(:,7);
G2=EXP2(:,8);
G2error=EXP2(:,9);
Y2=[X2;P2;G2];
t3=EXP3(:,1);
X3=EXP3(:,2);
X3error=EXP3(:,3);
N3=EXP3(:,4);
N3error=EXP3(:,5);
P3=EXP3(:,6);
P3error=EXP3(:,7);
G3=EXP3(:,8);
G3error=EXP3(:,9);
Y3=[X3;P3;G3];
t4=EXP4(:,1);
```

```
X4=EXP4(:,2);
X4error=EXP4(:,3);
N4=EXP4(:,4);
N4error=EXP4(:,5);
P4=EXP4(:,6);
P4error=EXP4(:,7);
G4=EXP4(:,8);
G4error=EXP4(:,9);
Y4=[X4;P4;G4];
t5=EXP5(:,1);
X5=EXP5(:,2);
X5error=EXP5(:,3);
N5=EXP5(:,4);
N5error=EXP5(:,5);
P5=EXP5(:,6);
P5error=EXP5(:,7);
G5=EXP5(:,8);
G5error=EXP5(:,9);
Y5=[X5;P5;G5];

Y=[Y1;Y2;Y3;Y4;Y5];
knobs=[[t1;N1(1);G1(1)];[t2;N2(1);G2(1)];[t3;N3(1);G3(1)];[t4;
N4(1);G4(1)];[t5;N5(1);G5(1)]]; %Dummy variable carrying
inputs needed for other functions

%Initial parameter guesses
g0=[1.8212 0.3377 0.2320 0.2168 24.7665 0.0023 0.4075 1.9382
0.0813 5.1013];

%Call nonlinear regression function
iter=statset('MaxIter',1);
[betahat,resid,J]=nlinfit(knobs,Y,@thefun,g0,iter);

%Display fitted parameters result
{'Parameter','Value';
'Yxn',[betahat(1)];
'Yxs',[betahat(2)];
'Yps',[betahat(3)];
'bmax',[betahat(4)];
'Kg',[betahat(5)];
'K',[betahat(6)];
'a',[betahat(7)];
'Yxp',[betahat(8)];
'umax',[betahat(9)];
'Kp',[betahat(10)]}

%Get confidence intervals for the estimated parameters
```

```
betahatCI=nlparci(betahat,resid,'jacobian',J)
flagmain=1;
return;
%=============================================================
============
function Ymodel2 = thefun(g,knobs)
Ymodel=0;
for nexp=1:5
    if nexp==1
        timevec=knobs(1:8);
        N0=knobs(9);
        G0=knobs(10);
    else if nexp==2
            timevec=knobs(11:18);
            N0=knobs(19);
            G0=knobs(20);
        else if nexp==3
                timevec=knobs(21:28);
                N0=knobs(29);
                G0=knobs(30);
            else if nexp==4
                    timevec=knobs(31:38);
                    N0=knobs(39);
                    G0=knobs(40);
                else if nexp==5
                        timevec=knobs(41:48);
                        N0=knobs(49);
                        G0=knobs(50);
                end
            end
        end
    end
end

%Initial concentrations (X, N, P, G)
Y0 = [0.1 N0 0 G0];
%Call ode solver
options=odeset('RelTol',1e-2,'AbsTol',1e-3);
[t,that]=ode15s(@(t,Y) simul(t,Y,g),timevec,Y0,options);
Ymodel=[Ymodel;[that(:,1);that(:,3);that(:,4)]];
end
Ymodel2=Ymodel(2:length(Ymodel));
return;

%=============================================================
============
function f = simul(t,Y,g)
```

```
%Define parameters for rate constant calculations
Yxn=g(1);
Yxs=g(2);
Yps=g(3);
umaxS=0.185;
bmax=g(4);
Kg=g(5);
K=g(6);
a=g(7);
Yxp=g(8);
umaxP=g(9);
Kp=g(10);

%Define species concentrations
X = Y(1);
N = Y(2);
P = Y(3);
G = Y(4);

%Define u and b
if (G>0.0001 && N>0.0001)
    u=(umaxS*N*G)/(K+(N*G));
elseif (G<0.0001 && N>0.0001)
    u=(umaxP*P*N)/(Kp+(P*N));
else
    u=0;
end
b=(bmax*G)/(Kg+G);

%X RESIDUAL BIOMASS TERM
rX = u*X;
%N NITROGEN CONSUMPTION TERM
if N>0.0001
    rN = -(u*X)/Yxn;
else
   rN=0;
   N=0;
end
%P PRODUCT FORMATION AND CONSUMPTION TERM
if (G>0.0001)
    rP = (a*u*X)+(b*X);
else
    rP=-u*X/Yxp;
end
%G GLUCOSE CONSUMPTION TERM
if G>0.0001
```

```
    rG = ((-a*u*X)/Yps)-((b*X)/Yps)-((u*X)/Yxs);
else
   rG=0;
   G=0;
end

f=zeros(4,1);
f(1)=rX;
f(2)=rN;
f(3)=rP;
f(4)=rG;

return;
```